DOCUMENT RESUME

ED 145 955 PS 009 633

AUTHOR Broughton, John
TITLE The Cognitive-Developmental Approach to Morality: A Reply to Kurtines and Greif.
PUB DATE Oct 75
NOTE 65p.

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS *Cognitive Development; Developmental Psychology; *Developmental Stages; Literature Reviews; *Moral Development; *Research; *Review (Reexamination)
IDENTIFIERS Greif (Esther Blank); *Kohlberg (Lawrence); Kurtines (William)

ABSTRACT

In this paper a recent critique of Kohlberg's cognitive-developmental approach to morality by Krutines and Greif is examined in detail. The critique is found to contribute some useful indications of needed study. However, a careful analysis of the literature on moral development shows the criticisms to be largely unfounded, in several cases misrepresentative of research findings. As a result, the empirical support for both "stage" and "sequence" components of Kohlberg's theory is somewhat underestimated. The misunderstandings reflect a common and mistaken attempt to reduce structural development in cognition to static psychometric traits, an attempt which confuses developmental theories with measurement instruments. (Author/SB)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED145955

Oct., 1975.

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

The Cognitive-Developmental Approach to Morality: A Reply to Kurtines and Greif

John Broughton

Wayne State University

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

John Broughton

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

PS 009633

Running head: The Developmental Approach to Morality

Abstract

A recent critique of Kohlberg's cognitive-developmental approach to morality by Kurtines and Greif is examined in detail. It is found to contribute some useful indications of needed study. However, a careful analysis of the literature on moral development shows the criticisms to be largely unfounded, in several cases misrepresentative of research findings. As a result the empirical support for both "stage" and "sequence" components of Kohlberg's theory is somewhat underestimated. The misunderstandings largely reflect a common and mistaken attempt to reduce structural development in cognition to static psychometric traits, an attempt which confuses developmental theories with measurement instruments.

Author's address: Department of Psychology, Wayne State University, Detroit, Michigan 48202

The Cognitive-Developmental Approach to Morality: A Reply to Kurtines and Greif[1]

William Kurtines and Esther Blank Greif (this journal, 1974, Vol. 81, 453-470) have brought our attention to some apparent weaknesses and anomalies in research on Kohlberg's account of moral development, and particularly to sins of omission on the part of Kohlberg himself. They correctly point to incompleteness in the data, especially with respect to the higher, or "post-conventional" stages. They also point to a widely-felt need for clarification of scoring procedures. This response attempts to show that most of Kurtines and Greif's other points are interesting queries deserving of an answer, rather than well-founded objections that might seriously undermine Kohlberg's claims. These queries turn out to be generated not only by lack of familiarity with the literature, but also by a broad misunderstanding of Kohlberg's theory. The conflict they sense is found to derive from a research paradigm and a conception of science and human mind quite at odds with Kohlberg's. This kind of difference has elsewhere been described as that between Aristotelian and Galileian modes of thoughts (Lewin, 1931; Elkind, 1969). In general, critical zeal in the application of this non-structural, non-developmental Aristotelian paradigm to a structural-developmental or Galileian theory appears to have encouraged pervasive misinterpretations of the research findings.

Specifically, Kurtines and Greif do not see that criticisms of

Kohlberg's measurement procedure do not necessarily imply criticisms of Kohlberg's theory. The theory could be correct, and we could still debate the best way to bring a measurement instrument to bear on it. The result of this misunderstanding is an ineffective criticism of a theory that needs and deserves a more thoughtful review, and which is susceptible to much more telling critiques. Kohlberg's case needs restating so that we are not distracted from these more fundamental problems.

More broadly, the misunderstanding reflects a not uncommon confusion about "Piagetian" theory and research, and the claims that they make. It is therefore instructive to take this opportunity to clarify such general confusions, a task which Kohlberg has not taken seriously enough until very recently (1975a, d). He could in fact be seen as contributing to this confusion by not making explicit his latent reconception of measurement, and by occasionally confusing his model with the more traditional psychometric one (Loevinger, 1974, pp. 59-60).

Focusing on the kinds of misinterpretations made by Kurtines and Greif, we should first draw attention to remarks seeming to indicate that Kohlberg has fulfilled neither of the two criteria most defining of a cognitive-developmental stage theory: "stage" and "sequence." Although in Kurtines and Greif's critique these issues are disguised as test "reliability" and "validity," they are better dealt with in Kohlberg's original terms, so that questions about the developmental theory are not assimilated to the measurement issues.

STAGES

Consistency

Stages are "structured wholes" (structures d'ensemble), characterized by their internally consistent organization of categories and relations (Piaget, 1960; Pinard & Laurendeau, 1969). In discussing within-subject distributions of reasoning across different stages, (p. 457) Kurtines and Greif state quite bluntly that "a majority of an individual's responses are rarely at his modal stage." On the same page, apropos of inter-item or inter-story homogeneity, they say that "there are no estimates of internal consistency to be found anywhere in the research literature." Although there are two intertwined issues here (self-consistency of moral thought, and test homogeneity of moral situations), Kurtines and Greif's position on both appears to be directly in conflict with this passage from Kohlberg's major treatise on moral development theory:

> In addition to sequence, stages must meet the criterion of consistency implied by the notion of a "structured whole." On the logical side, consistency is found in the fact that twenty-five distinct aspects of moral judgment may be logically defined from the core concepts of the six stages. On the empirical side, both consistency across aspects and consistency across verbal situations are to be found. Such consistency is to be indicated, first, by the fact that an average of 50 per cent of a subject's moral judgments fit a single stage.[2] Second, such consistency is indicated by fairly high correlations in moral levels from

> one story to another. The highest such correlation is .75, the lowest .31, the median .51. Third, it is indicated by the fact that these correlations between situations are not specific, i.e., there is a general moral level factor. A general first moral level factor accounts for most of covariations from situation to situation. (Kohlberg, 1969a, pp. 388-9)[3 4]

A more thorough factor analysis of story-item scores has been carried out by Lieberman (1969).[5] Although a median correlation of .51 means that only about 25% of the variance is accounted for, there is other support for Kohlberg's original consistency claim. Haan et al. (1968) found that over half of 957 subjects had a percentage of reasoning at their modal stage that was at least twice as great as the percentage found at any other stage. In Turiel's study (1966), the proportion of subjects with a mode this dominant was two thirds. Also, Turiel (1969) and Rest et al. (1969) reported that, on average, the modal percentage is two and a half times as great as the percentage use of either adjacent stage. Finally, we have Kohlberg's own "quasi-simplex" data (1963a), which shows that only adjacent stages ever have positive correlations with each other. Of course, our interpretation of these findings is jaundiced by the indications that scorer bias might be partly responsible for the apparent consistency.

Underlying much of this confusion in the review is a simple oversight, aggravated by delayed publications of Kohlberg et al. Kurtines and Greif base their criticisms on Kohlberg's 1958 dissertation, in which consistency is indeed reported as considerably lower than the

50% mark. The 1958 scoring system has been widely used, often in a form elaborated by Porter and Taylor (1972). However, that rather crude measure was abandoned six years ago in favour of a more "structural" measure, (Kohlberg, 1969c[6]), and it was this new measure that was used in several recent studies to which Kurtines and Greif refer (e.g., Holstein, 1972; Saltzstein, Diamond, & Belenky, 1972). "Structural" again means simply that the measure (like Piaget's logical tasks) taps the form of thought, or the way in which reasoning is organised, rather than the specific content of the response, a distinction which Kohlberg's critics do not take into account. In line with this reorientation, there was also a change in the units across which consistency is assessed. In the new method, subjects are given a score on each of nine moral "issues," rather than on each story-item. These issues comprise the set of basic moral topics to which moral reasoning orients (rights, duties, punishment, etc.). They are derived from the theory, thereby contributing more to the construct validity of the measure. Lieberman (personal communication) has also carried out a factor analysis of subject scores on these nine moral issues of the 1969 "structural measure," to establish the consistency within subjects over issues. Each issue was scored by an independent scorer, thus eliminating halo effects (a precaution not taken seriously enough in previous assessments of Kohlberg's instrument). A first-level factor accounted for over 85% of the variance, which is more than three times as much as the variance accounted for in Kohlberg's original study (see the above quotation).[7]

In summary then, Kurtines and Greif's remarks on consistency need to be seen in the light of the fact that they bear on an obsolete measure, that there was a certain amount of within-subjects consistency even with that old measure, and finally that there is greatly improved consistency across moral issues using the new "structural measure." An informed, critical and detailed history of the Kohlberg system is reported by Loevinger (in press).[8]

Reliability

Kurtines and Greif are correct in raising the question of test-retest reliability, since Kohlberg has not been very vocal on the topic. However, they make the puzzling claim that "evidence from experimental studies indicates that scores on the scale fluctuate greatly even over short periods of time," (p. 468). The claim is puzzling, since on p. 465 the authors seem to be arguing that because Turiel was unable to obtain significant improvements in Kohlberg's score through training, we should conservatively treat the results as showing no change. There appears to be some conflict here, since either Turiel's results support Turiel's thesis, or else they at least provide us incidental evidence of temporal stability. The issue of stability of reasoning over time is of course very important to the notion of structural stages, being in fact one of the main criteria of structural cognition (Smedslund, 1961), but Kurtines and Greif do not indicate when they are talking about structural stability and when they are talking about measurement stability. Despite this, they do a service in pointing to inadequate treatment of test/retest

reliability in cognitive-developmental literature. The only data we have been offered on this (Kohlberg, 1975c, Part II, p. 84) is with an N of 20, rated by Kohlberg himself (and hence presumably not done blind), which yields the exact same test and retest stage score in 17 out of 20 cases. The time gap involved is not reported.

Kurtines and Greif devote little space to the important form of reliability called inter-judge. On the same page, Kohlberg tells us the degree of exact agreement is variously 70% (Hickey, 1972) and 86% (Saltzstein et al., 1972), the N being again 20 in both instances. Exactly what inter-rater reliability means is not discussed. On this count, and many others, Kohlberg can be questioned for too uncritical an acceptance of standard social science methodology (cf. Morelli, 1975).

To conclude this section, while Kurtines and Greif point to interesting lacunae where more empirical work would be informative, they also confuse issues by failing to grasp the significance of the claim that moral judgments form a coherent and consistent network called a structural stage. This leads to a failure to distinguish measurement issues from developmental-theoretical ones (particularly confusing reliability with structural stability as well as structural wholeness), and to a failure to distinguish imperfections in scoring criteria from disconfirmations of empirical claims. Kohlberg has contributed to this problem by assuming these distinctions, rather than explicating them (Loevinger, 1974, pp. 59-60), and by a confusing use of techniques like the Guttman quasi-simplex. While

Kurtines and Greif's psychometric criticisms fail to be appropriate to a developmental stage sequence that is not just a psychometric scale, were those criticisms factually adequate and lodged in the context of a stage-structural theory they might be of more interest. As it is they really only call for a slightly more explicit account of matters from Kohlberg--which in fact is needed.

SEQUENCE

In order to show that Kurtines and Greif's discussion does not strongly challenge Kohlberg's claims for stages in moral development, we were obliged to focus several partial and often tangential suggestions in that discussion which were in danger of being seen as entirely disputes about scoring methods or reliability. When we come to consider the other major criterion of cognitive-developmental stage theories, that they describe invariant sequences, we have less focussing to do since the arguments are more explicit. But we still find the assimilation of Kohlberg's theory to an inappropriate set of evaluative criteria, because questions of stage "sequence" are addressed as though they were only questions of psychometric construct validity. It is not at all clear why the matter of sequence is considered to be an issue of construct validity, while the matter of stages is not. This preferential treatment seems to be quite at variance with Kohlberg's own claims in his "Stage and Sequence" manifesto, (Kohlberg, 1969a), where he treats the two as correlative theoretical issues.

Cross-cultural evidence

Kurtines and Greif are quite correct in pointing to the limitations of age trend data as support for claims about sequence. Kohlberg has pointed this out himself (1958, 1969a, 1973a) and clearly emphasizes that only longitudinal data can substantiate invariant sequential properties. This is one reason for the considerable publication delays involved in authentic cognitive-developmental research. It is a criticism that bears repeating (cf. also Costa, 1971), because age trends have often been the only support offered for a developmental theory (e.g., Perry, 1968).

However, when we come to the remark that the demonstrated recurrence of moral stages across cultures provides "no support for the claim that the stages are universal" (p. 462) we are left a little puzzled, since if Kurtines and Greif mean to refer to problems of induction (cf. Kaplan, 1967; Zinn, 1970; Morelli, 1975), this cryptic statement needs some expansion. Although necessarily inconclusive, and irrelevant to the issue of sequence, it seems that such cross-cultural data are quite germane to the issue of universality and provide an empirical test of Kohlberg's claims. A recent, highly critical review of Kohlberg's universality claims by Simpson even acknowledges this evidence as "suggestive," (Simpson, 1974, p. 86). Of course, what is needed again is longitudinal data which we are led to believe are forthcoming (Turiel, Kohlberg, & Edwards, in press), and which should bear on the issue of universality of sequence as well as the universality of the stages. Needless to say, it is not only Kurtines and

Greif who eagerly await these findings.

The quasi-simplex correlation matrix

Kohlberg's second major form of evidence for the sequential properties of stages is the Guttman "quasi-simplex" matrix of correlations. This kind of evidence is no longer necessary, since it is clearly pre-empted by longitudinal studies of sequence. However, it has considerably more validity than Kurtines and Greif admit. They say that the quasi-simplex "provides no direct support for the invariance of the six stages." Kohlberg has in a sense already admitted this when he says that the pattern of correlations is only compatible with a cognitive-developmental interpretation. But Kurtines and Greif's objection is that if we reverse the stages within levels of moral reasoning, (exchanging columns 1 and 2, 3 and 4, and 5 and 6), that the matrix of correlations shows a very similar quasi-simplex pattern to the ordered matrix.

First, we might ask why the stages should be jumbled in this way. If Kurtines and Greif want to make a strong point against invariant sequence, why do they not simply randomly mix up the columns? The reason is presumably because there is some order here, perhaps the order between levels if not the order between stages. Thus their choice of transformation to apply to the matrix implicitly supports a weaker version of Kohlberg's claims. Kohlberg has pointed out anyway that the major transitions are between levels (1963a, p. 17).

Second, we might ask exactly what Kurtines and Greif expect us to see as "similar" between the two matrices. The quasi-simplex

pattern (correlations decreasing down columns) is violated by only two pairs of correlations (out of ten) in Kohlberg's matrix of actual data, whereas it is violated by five pairs in Kurtines and Greif's transformation.

Finally, it is a little unreasonable of Kurtines and Greif to chide Kohlberg for the lack of independence in the matrix, since the latter made this circularity quite clear:

> This fit to our schema ... is that of intuitive conception of empirical relations in a body of data to a systematic quantitation of these relations in the same data. (Kohlberg, 1958, p. 101)

Despite other criticisms of this kind (Costa, 1971), it should be clear that ideal-typical constructions necessarily proceed in this reflexive fashion (Weber, 1949; van den Daele, 1968). The construction of ideal types is a fusion of analytic and synthetic meaning, and cannot conform to the traditional hypothesis-testing paradigm of normal science, since that requires strict segregation of a priori hypothesis and a posteriori test. This leaves us, however, with the persisting problem of how to reconceptualize the process of confirmation and disconfirmation when we are dealing with ideal types (Rex, 1975). Kohlberg has not done justice to this issue either.

Longitudinal evidence

The third and most important form of evidence that Kohlberg adduces in support of the sequential ordering of his moral stages is from longitudinal research. Kohlberg's longitudinal findings are

only just now appearing (Kohlberg & Elfinbein, in press, table 3). Of 30 subjects followed at 3 year intervals from age 10 to somewhere between ages 19 and 28, all are reported as either remaining at the same stage, or proceeding as much as 3 stages through the prescribed sequence. In only 2 cases was there more than one stage advance in the interval between testings. Development appears to be gradual and stage by stage, transition from one to the next taking 4-5 years on average in adolescence, and tending (with some exceptions) to reach an apparent ceiling around age 20. Interesting and important as this evidence is, scoring of the longitudinal data was not uniformly blind with respect to age, and so vital sources of scorer bias have still to be dealt with before this evidence can be properly evaluated. Loevinger (1973) is a good example of someone taking this kind of problem seriously and dealing with it (cf. also Broughton, 1974, 1975).

An earlier and controversial report of some specific longitudinal findings is Kohlberg and Kramer's (1969). Kurtines and Greif rightly point to the anomalous "sophomore regression," found in this longitudinal follow-up of Kohlberg's original subjects, as a violation of strict stage-by-stage development. Kohlberg and Kramer originally interpreted the recurrence of egoistic stage 2 thought in the transition from stage 4 to stage 5 as Wernerian "spirality" or Kris's "regression in the service of the ego," being of ultimately progressive nature.

This does not necessarily imply that the stage sequence is

"flexible" as Kramer (1968) originally hypothesized, and as Kurtines and Greif conclude, since all the "regressors" were eventually found to advance to stage 5. It is therefore unjust of Kurtines and Greif to say (p. 469) that "longitudinal data demonstrated only that most males are at stage 4 morality." Furthermore, a detailed analysis of the actual profiles of these subjects shows that the "return to stage 2 reasoning" was only a partial one (Broughton, 1970).

Fortunately, however, we do not have to debate this issue, since the "regression" turns out to be no regression at all! With a more adequate differentiation of form and content in the interpretation of late adolescent moral judgment, what was classified "regressive stage 2 thinking" emerges as a new transitional form of judgment, colloquially referred to as "stage 4 1/2" (Kohlberg, 1973a; Gilligan & Kohlberg, 1974). Turiel (1974a) has argued that this is neither a stage nor a substage but simply reflects a transitional reformulation of stage 4 into stage 5. There is, however, general agreement between Turiel and Kohlberg that the moral outlook at this stage is egoistic in a new sense, bearing only a superficial similarity of content to the hedonism of stage 2; it is a sophisticated meta-ethical individualism (Boyd, 1972), presupposing ethical relativism:

> the assertion that a "good" moral reason, rule or principle was one that appeals to the actor's self-interest or to his point of view, where the actor's individual point of view is "natural," "outside society," "amoral," or independent of acceptance of the rules and expectations of others. (Kohlberg, 1973a, p. 192)

Such an ethical position is clearly post-conventional, and not a form of naive preconventional hedonism like stage 2. It was only poor probing of the original interview responses that permitted the occurrence of such a confusion of superficial content with deeper formal qualities of reasoning.[9]

The removal of Kurtines and Greif's objection to "flexibility" in the stage sequence clearly extends to their remarks about the Haan et al. (1968) study of political activism in the Berkeley Free Speech Movement sit-in. Kurtines and Greif argue that, since stage 2 and stage 6 individuals were both civilly disobedient, this raised questions about the discriminant validity of the stages. In the face of the recent reinterpretation of "regressed" moral reasoning, it now seems questionable to claim that this study shows that "the actual behavior of Stage 2 and Stage 6 males was similar." (Kurtines & Greif, p. 459). An alternative interpretation of the Haan study is presented below, under the rubric of "moral action" research.

Intervention studies

Although Kurtines and Greif approach the fourth form of evidence for sequence, accelerated transitions between stages, they confine their attention to a single study - that of Turiel (1966, expanded in Turiel, 1969, 1972). The sort of evidence obtained from this kind of study is particularly important to the stage theory, for several reasons. First, there is the long waiting period for both research and publication in longitudinal studies of development. Second, there is the fact that developmental theories have been wont to stress the

static, stage aspects of development rather than its dynamic transition aspects (Wohlwill, 1970; Turiel, 1969; Riegel, 1974; Chomsky, 1968). Third, if we are to accurately characterize development, we must be able to characterize the particular constellations of experience that lead to developmental advance (Wohlwill, 1973b; Zinn, 1970). Studies of transition test the explanations of the equilibration of thought that cognitive-developmental theories offer us. Fourth, it is a strong test of the sequence claims of the theory if we can show that intensified experience with, say, moral conflicts produces movement towards the next highest stage, however small.

It is therefore important to realize that there is additional evidence concerning accelerated transition, evidence which spans the whole range of stages. Blatt (1975) arranged classroom moral discussions in which judgments were constantly juxtaposed with arguments at the next highest stage. He found that reasoning at stages 2, 3, and 4 regularly tended to move toward stages 3, 4 and 5 respectively. Hickey (1972) obtained similar results with the lower stages, through guided moral discussion in a prison setting. Erdynast (1973) obtained partial movements from stage 5 towards stage 6 in business executives under guided instruction (see also Kohlberg, 1973b). Another study in progress (cf. Kohlberg, 1973b) has found educationally induced transition, occurring in the direction specified by the theory of moral development.[10] Kohlbergian studies in education are reviewed by Rest (1974), and more cautiously (although less comprehensively) in a good treatment by Brown and Herrnstein (1974). A thorough

evaluation of Kohlberg's claims for invariant sequence must also cover these studies.

In summarizing our findings so far, while the data is clearly incomplete, especially in regard to the higher stages, we see considerable cause to question Kurtines and Greif's criticisms both insofar as they bear on the stage-like consistency of individual moral reasoning, and on the sequential ordering of these stages.

Turiel's Study

The final source relevant to sequence considered by Kurtines and Greif is Turiel's important study, in which he tried to experimentally induce small upward and downward changes in stage of moral reasoning.[11] Kurtines and Greif attempt to undermine Turiel's basic conclusion that the number of +1 scores of the +1 treatment group was higher than the number of +1 scores of any other treatment group. First, (p. 464) they argue in a bizarre fashion that this claim "is at best ambiguous however, since the actual ts are unreported," even though the significance levels are! (For the curious, the t values are reported in the original dissertation.) Second, they state that "Turiel claimed that the number of -1 scores of the -1 treatment group was not significantly higher than the -1 scores of any other treatment group." This is not really true. Turiel (p. 616) did conclude that the -1 treatment had an effect in moving subjects down one stage. Kurtines and Greif arrived at this false impression by failing to notice that Turiel's t-test was a Dunnett t-test for the simultaneous comparison of one group mean with several others (Winer,

1962). As a result, they concluded wrongly that the difference was greater than the borderline significant difference between +1 scores for the +1 group and -1 scores for the -1 group. In fact, when we take into account the different relation of t values and significance levels for the Dunnett, the former difference turns out to be less than the latter, supporting Turiel.

In addition, Kurtines and Greif make the following major points: i) change scores are very small; ii) the difference between +1 and -1 change scores is not significant at the standard 0.05 level; iii) whatever effects are observed are attributable to memory rather than to structural change; and iv) instability in the moral judgment score renders any change measures uninterpretible.

i) It is true that the changes are small, but this should come as no surprise if we accept the interactional basis of the structural explanation of developmental advance. The amount of logical conflict generated in the experimental situation must be minimal compared to the kind and enormity of conflicts that arise in the repeated everyday situations that are morally challenging. The changes in direct scores are only small relative to empiricist expectations that moral reasoning is impressed upon people through a process of social coercion. Turiel's loss is in this respect the rationalist's gain, since large changes would have been an indication of the predominance of accommodatory over assimilatory tendencies in moral judgment. The developmental interpretation is also supported by the very limited generalization reflected in the indirect scores.

ii) Kurtines and Greif's emphasis on strict adherence to significance levels is somewhat regressive in the face of the powerful critiques of such behavior over the last 15 years (e.g., Levine, 1974; Cohen, 1962; and Bakan, 1966). The point is that changes in the +1 group were one and a half times as great as changes in the -1 group, and so it is not strictly true for Kurtines and Greif to say that "both seem equally effective" (p. 464). The existence of an effect, and its possible theoretical significance are warrant enough for holding judgment in abeyance here. The significance level does not measure the size of the effect. Kurtines and Greif are quick to run the risk of type II error, even though they present no argument that we should guard more carefully against type I error in this study (Cohen, 1966). At another point Kurtines and Greif switch horses, as we mentioned already, arguing that the general effectiveness of the +1 training is not demonstrable, despite a significance level of $p < .005$, simply because the actual t value is not given!

iii) The objection that change effects are due to memory is directly in conflict with the pattern of results we have just discussed.

iv) Since Turiel subtracts the appropriate control group changes from the measured effects in the experimental groups, "instability" in the moral judgment score is already taken into account in the results. Furthermore, Turiel does justify this subtraction procedure, despite Kurtines and Greif's allegations to the contrary (p. 464). It is not

clear anyway what Kurtines and Greif's objections to this procedure could be. It makes more sense simply to inquire how the spontaneous regression is to be explained (Zinn, 1970). My experience with repeated administration of the Kohlberg interview suggests that declining interest in the test leads to shorter and less careful answers, which in turn generate artifactual lowering of ascribed scores.

In sum, although Kurtines and Greif are correct to point to a certain inconclusiveness in Turiel's results, their arguments are often misleading, and their statistical dogmatism obscures important information that may require explanation and further study. The critique seems particularly unjust, not solely because of its distortions, but because it isolates issues from the context of the study as a whole, and from a body of research which is consonant with Turiel's findings. For example, Rest's finding that 12 and 16 year-olds reject reasoning at the -1 level, and the analyses of change in studies already mentioned (e.g., Hickey, 1972), confirm Turiel's claims. Furthermore, Keasey (1973) and Tracey and Cross (1973) have succeeded in replicating Turiel's findings. A study by Sutton (1972), cited by Turiel (1974a), adds further support, despite starting out with radically different hypotheses. The only ambiguous findings are those of Arbuthnot (1975) which confirm only some of Turiel's hypotheses. Unfortunately, the apparent disconfirmation is not documented in Arbuthnot's paper. A fairer and more incisive critique of Turiel's work is presented in Zinn (1970),

who points to the fact that the "+1/-1" nomenclature is based on the false premise of equal interval scaling of moral stages, and that it prevents the analysis of changes in the whole profile of moral reasoning.

In summarizing our findings so far, we see considerable cause to question Kurtines and Greif's criticisms, both insofar as they bear on the stage-like consistency of individual moral reasoning, and on the sequential ordering of these stages.

RELATION OF MORAL JUDGMENT AND MORAL ACTION

In general, much as Kurtines and Greif overestimate the evidence against "stage" and "Sequence" claims, it seems that they underestimate the amount of empirical confirmation for the moral stages, and in particular the strong evidence of the relation of moral judgment to action. For example, they do not mention the fact that Kohlberg's stages are, to date, the best predictors found for the behavior of the subjects in Milgram's "shocking" experiment. Only subjects with principled moral reasoning questioned authority and quit the task (Kohlberg, 1970, 1969a, pp. 395-6; Milgram, 1974, p. 205). Moral stage predicted behavior in this situation while standard measures of sympathy did not. Kurtines and Greif also selectively exclude the discriminant validity of the moral stage definitions in situations where cheating behavior is assessed (Krebs, 1967). Kohlberg and Freundlich (in press) have found moral stage to predict recidivism in a study of post-release adjustment in delinquents. Finally, to counter Kurtines and Greif's suggestion that the research literature

is insufficiently independent of Kohlberg's influence, an independent study by Krebs and Rosenwald (1973), which set out with some scepticism towards Kohlberg's theory, found that moral stage predicted with surprising discrimination subjects' fulfillment of a commitment to return questionnaires. It seems therefore that we are justified in questioning Kurtines and Greif's conclusion (p. 468) that "there is no clearly demonstrated connection between moral judgment, as measured by the Moral Judgment Scale, and moral action."

Perhaps the most interesting research on moral judgment and action is the Haan et al. study of political activism in the Berkeley Free Speech Movement. In the light of Kohlberg's eaten words about "college repression," we promised above a reinterpretation of this study. A careful reading of the paper reveals support for Kohlberg's position rather than the disconfirmation that Kurtines and Greif claim. The latter argue that subjects scored stage 2 manifested the exact same protest behavior as subjects scored stage 6, thus compromising the discriminant validity of the theory.

First, a close examination of the data shows that Haan et al.'s "Stage 2" subjects are misclassified. When these subjects were questioned about their civil disobedience, they rejected the notion of reciprocal obligations, and their:

> reasons were more often concerned with their individual rights in a conflict of power. They frequently included the indubitable fact that the sit-in had worked to bring social-political reforms on the campus. (Haan et al., p. 198)

Notions of individual rights and of social-political reform are far beyond the conceptual grasp of stage 2 moral reasoning. They are developmental achievements reserved for stages 4 and 5, because it is only at those stages that the net of mutual relationships constituting the moral structure of society is finally conceptualized, and only within this conceptualization that reciprocal obligations could possibly be rejected. For real stage 2 thought, there are no such things to reject. Our argument is well supported by the fact that Haan et al. report their "stage 2" individuals' average profile of moral judgment stages (p. 187), and in that profile, stage 5 reasoning is the next most frequent to "stage 2" reasoning. Individuals who were really functioning chiefly at stage 2 would show some reasoning at stage 3 and some at stage 1, but none at stage 5. Instead we see something approaching a bimodal distribution, with the main peak at stage 2, and a second smaller peak at stage 5. The data suggests therefore that these subjects would fall under Turiel's description of the "4/5 transition" (cf. Sullivan & Quarter, 1972).

A second weakness to Kurtines and Greif's argument about discriminant validity is that it is difficult to see how the activism of "Stage 2 males" (however we reclassify them) could in any way be described as <u>the same behavior</u> as that of "Stage 6 males," when its orientation is pragmatic, goal-oriented confrontation, rather than the taking of a moral stance on the principles of justice underlying civil liberties. These two behaviors are worlds apart. To try to argue that in some concrete, spatial sense their "actual behavior

... was similar" is to resort to a primitive behaviorism that equates action with physical movement. Such a viewpoint, exemplified also by Sanders and Cone (1975), fails to acknowledge that it is impossible to define or describe a behavior unambiguously without reference to underlying reasons or intentions (cf. Merleau-Ponty, 1963; Hamlyn, 1970; Taylor, 1964; Chomsky, 1959). In fact, Haan et al. themselves are at pains to make clear the fundamental differences between the two types of activism, and to criticize a superficial equating of them (pp. 196-8).

Kurtines and Greif's disappointment about discriminant validity perhaps stems from the fact that Kohlberg has not supplied one-to-one behavioral accounts of his stages, i.e., he has not specified sets of six mutually exclusive behavioral alternatives in each of a range of morally conflictful situations. If this is the kind of "discriminant validity" that Kurtines and Greif are seeking, then it is an unjust request, because Kohlberg's stages describe <u>forms</u> of moral reasoning that can be used in support or rejection of a particular action. In addition, these forms are <u>logically interconnected</u> in a hierarchy of increasing moral adequacy; they are not pass/fail items. As Beilin (1971, p. 189) has pointed out, cognitive-developmental accounts stand and fall "by the explanatory power of the entire theory." By demanding separate behavioral confirmations for each stage, Kurtines and Greif are attempting to "dimensionalize" Kohlberg's theory (Costa, 1971), to impose the paradigm of a psychometric scale of discrete items upon a sequence of qualitatively different stages that are

neither developmentally discrete nor constitutive of a psychometric scale in the traditional sense.[12]

But the most fundamental point of Kohlberg's argument which this criticism misses is that the stages of moral development do not depend upon prediction of action for their validity, since a) these are stages of moral judgment, and b) there is no coherent separate domain of "moral behavior." Kohlberg's position is that moral behavior does not develop (Kohlberg, 1963b), nor could it as long as it is defined behavioristically apart from underlying reasoning (Kohlberg, 1975a). Correlatively, empirical findings show how situationally-specific moral behavior is. What Kurtines and Greif miss is that the relation between moral judgment and action within a cognitive-structural theory is quite unlike the relation traditionally described between traits of moral attitude and moral behavior. The structures of moral judgment determine action by defining how rights and duties are to be perceived in concrete situations (Kohlberg, 1969b). They are not "values" predicting to greater or less incidence of pro-social behavior. The only way in which a critique of the cognitive-developmental approach might have impact in this regard (see for example Westermarck, 1932, or Sanders and Cone, 1975), would be to call into question the basic assumption that morality involves intelligent judgment about rights and obligations (judgment which is necessary but not sufficient for moral action), and that the study of moral conduct therefore presupposes investigation of the reasoning processes underlying such judgments. However, as pointed out by Smythe (1974), even

if this investigation were completed, we would still be left with a rather poor understanding of exactly how reason and conduct are to be integrated conceptually. It seems that Kohlberg, unlike Dewey (1922) say, has himself brushed over the distinction between reasoning and action, and has tended to take Socrates' position that he who knows what is right will do what is right. This involves underestimating the dark side of human nature. Although there has been undue emphasis on "action for action's sake" (Horkheimer, 1974), empirical studies of concrete moral decisions are needed to elucidate the nature of responsible moral action.

"COUNTEREVIDENCE"

Under this heading Kurtines and Greif offer certain evidence meant to disconfirm Kohlberg's theory, although their first line of evidence bears on Piaget's stages of moral development, not Kohlberg's. Their main thrust here is to report studies reputedly showing that "moral judgments are directly affected by social influence." Now this in itself does not run counter to Kohlberg's theory at all, since Kohlberg has explicitly stated that development proceeds as a function of "role-taking opportunities," which is social experience par excellence. Kurtines and Greif are therefore guilty of a non sequitur when they say:

> modes of moral thought are subject to social influence; consequently, one would not necessarily expect them to follow a preset order. (p. 467)

It would be quite justifiable within Kohlberg's theoretical framework

to say that it is precisely because modes of moral thought are subject to social influence that they follow a preset order (if not a predetermined one), since it is specifically social experience that leads to the disequilibration of a person's system of moral reasoning, and which motivates re-equilibration at the next highest stage (Turiel, 1974b).

Modeling studies

What is correct about Kurtines and Greif's intended argument is the much more specific point that a developmental theory minimizes the explanatory power of contingent or extrinsically imposed social conditioning. In operational terms, this implies that reversible changes in moral judgment under the control of experimenter or model are evidence against the cognitive-developmental thesis. Kohlberg denies, not that moral judgment is influenced socially, but that it is totally the product of social learning, or copying. Whereas Kurtines and Greif take Bandura and McDonald's data to imply this, Kohlberg and others do not. For example, Kurtines and Greif say:

> While interpretations of the meaning of these results have been variable (cf. Cowan et al., 1969), the implications of the data are clear - modes of moral thought are subject to social influence;

But even though they replicate the findings, Cowan et al. do not see the implications as clearly as Kurtines and Greif would like to think. Even in their abstract we find:

> A theoretical analysis showed that neither the present study nor

> that of Bandura and McDonald could be used directly to affirm or deny Piaget's hypothesis ... Some of the findings raise questions which cannot yet be answered within the social learning framework. (Cowan et al., p. 261)

The implications of Bandura and McDonald's study are therefore not unequivocal. Cowan et al. object to the fact that atypical subjects were used who were already transitional between stages, that inferences about development were drawn from group means, and that choice data only, rather than explanations, were elicited in the post-test (although Bandura, 1969, appears to dispute this latter criticism). One of the findings which "raise questions" was the fact that subjects who were conditioned "up" showed much more pronounced modeling effects in the two-week post-test than did subjects that were conditioned "down." This result does not fit with the social learning notion that moral responses are simply habits, but it does fit well with the cognitive-developmental theory of unidirectional progress, and the related research findings in attempts to accelerate moral development (Turiel, 1966, 1969). Another finding with similar implications is that there was more change to the lower level by the higher level subjects than vice versa, a finding uninterpretible from a social learning point of view, but quite consonant with Kohlberg's theory, and with cognitive-developmental research on the relative difficulty of using reasoning at a level ahead of your present stage (Rest et al., 1969). Additional evidence from Dworkin (1968) and LeFurgy and Woloshin (1969) suggests that social learning effects vary considerably

with the subject's developmental level (Zinn, 1960).

A second problem for the social learning account is explaining the existence in children of moral stages lower than adult stages (Cowan et al., p. 263). How can such forms of moral thought be learned through imitation if there are no models responding or instructing in this way? Bandura (1969) suggests that adult morality is inconsistent, and therefore often does provide a model, and reinforcement, for lower level response. Although Bandura adduces no evidence to support it, this argument seems prima facie reasonable. Although research evidence indicates that adult moral reasoning normally operates quite consistently at a higher level (Kohlberg, 1969a; Turiel, 1969), parents do appear to adjust the level of their discourse to the perceived intellectual capacity of the child. However, there is also evidence from the Rest et al. (1969) study that this kind of downward adjustment involves a change in content only and not in form of moral judgment.

A third problem for the social learning account is that when the moral has been viewed as behavioral conformity in empirical research (e.g., Turner, 1952; Hartshorne & May, 1922), no age increases in such moral traits as altruism or honesty have been found (Kohlberg, 1963b). In fact, Turiel (1974b) has recently shown that knowledge of cultural values decreases with age. Simpson (1974) has attempted to negate this argument of Kohlberg's by pointing out that honesty and altruism may only be espoused not enacted social ideals. However, if this is the case, then there should be increasing

conformity to these "actual" traits with age, and it is evidence of any such increase that is notably lacking.

What gets completely concealed in Kurtines and Greif's account is that the whole debate is over the stages in Piaget's theory of moral development, not Kohlberg's. It is not even clear that Kurtines and Greif are aware of this. Kohlberg (1969a, pp. 404-9) has stressed that a test of Piaget's moral theory is not a test of cognitive-developmental theories of morality in general, since Piaget's behavior definitions do not have any structural component anyway. In other words, Piaget's moral stages (unlike his logical ones) do not satisfy his own criteria for true structural stages (Kohlberg, 1963b). The implications of this for the interpretation of Bandura and McDonald are as follows:

> Piaget's definitions of moral judgment responses do not meet the naturalistic criteria of structural stages, and hence social-learning manipulations of these responses do not indicate that moral judgment is not structural, nor do they indicate that social-learning operations can account for structural development. (Kohlberg, 1969a, p. 408)

Turiel (1966), Kohlberg, and Cowan et al. point out that what Bandura and McDonald have demonstrated in concentrating on the reinforcement of moral response, is the learning of the expected answer, under social coercion. Kuhn and Langer (1968) have demonstrated that the usual modeling experiment conveys implicit instructions to the child to do what the model does. Bandura and McDonald could not have

demonstrated structural change as a result of their manipulations, since the dimension manipulated ("subjective versus objective responsibility") is but one of eleven observable aspects of the structure of Piaget's stages (Kohlberg, 1963b; Turiel, 1969). Also, to demonstrate structural change, one must demonstrate stability of reasoning over time (Smedslund, 1961). Instead, Bandura and McDonald found that learned moral responses were readily reversible.

We can conclude that reversible changes in moral judgment under the control of experimenter and model have yet to be demonstrated. However, Kohlberg will continue to be vulnerable to objections of this ilk until he specifies exactly what the nature of moral experience is, and the precise relation of cognitive structures to his vaguely characterized "structures of the environment" (Riegel, 1975b; Sanders & Cone, 1975).

Sex differences

A second line of "counterevidence" argues that, in a study by Holstein (1972), "cross-sectional sex differences in modal response did not support the notion of an invariant sequence" (Kurtines & Greif, p. 467). Kurtines and Greif's argument is simply that the modal stage of reasoning for males was different from the modal stage for females. How this bears on the invariance of the sequence at all is a complete mystery. Differences in rate of development are expected, indeed predicted by the theory (Kohlberg, 1958; Smith & Firestone, 1973; Kohlberg & Gilligan, 1971). But neither differences in rate nor differences in ceiling of development challenge

the sequential nature of the stages.[13]

Perhaps there is a confusion here with a different study of Holstein's (1973) in which she argues that because 23% of an adolescent sample were two stages higher when retested after three years, they must have skipped a stage (cf. also Smith & Firestone, 1973). But a more obvious interpretation is that these individuals happened to pass through two consecutive stages in the period of three years (Turiel, personal communication; cf. data in Kohlberg & Elfinbein).

Personality correlates

A third line of "counterevidence" appears to be as ineffective as the second one. Hogan (1973) is meant to have developed a scale that distinguishes "utilitarian, legal" versus "intuitive, principled" thinking. These types of thinking are found to be associated with particular personality characteristics. Finally, we are told, "the more mature individuals tend to cluster in the center of this dimension." We are supposed to infer that stages 5 and 6 are therefore shown to be "alternative, but equally valid, forms of moral thought."

First of all we might ask whether non-moral personality differences are strictly relevant to the structure of moral thought. Despite Hogan's claims (1974), Kohlberg has clearly shown that morality is not reducible to personality dimensions such as more or less autonomy, more or less intuitiveness, or more or less rule-orientedness (Kohlberg, 1969a, 1971). He has specifically militated against the artificial opposition of rule-oriented and principled morality:

> such a dichotomy between conformist morality and rational

morality is fatal for the study of both. (Kohlberg, 1958, p. 4) Nor can there be "too much morality" (Hogan, 1974, p. 116) since morality is not a quantitative matter (Kohlberg, 1958, p. 7). In addition, Hogan certainly does not show that "personality differences are important determinants of an individual's style of moral reasoning" (Kurtines & Greif, p. 469). He seeks only to research correlates (Hogan, 1973, p. 225).

Putting these problems aside for the moment, we should note that Hogan asserted an ideological opposition to Kohlberg's view of morality, he did not seek an empirical test of it, or present evidence at variance with Kohlberg's theory. The explicit alignment of Hogan's moral traits or "dispositions" with Kohlberg's stages 5 and 6 appears to belong to Kurtines and Greif. In so doing, they in fact misinterpret Kohlberg, confusing stage 5 with stage 4, and stage 6 with a version of "stage 4 1/2." Stage 5 reasoning does not imply conventionality and resistance to change, as Kurtines and Greif report. Nor is it the stage of moral reasoning characteristic of policemen or ROTC seniors. Neither can stage 6 reasoning in any way incur an association with impulsiveness, opportunism, and irresponsibility. And its focus is not necessarily civil disobedience.

Such caricatures reflect the improper a priori dichotomy we have just mentioned. As Kohlberg notes:

> If the basic concept of morality is conformity, then all the more subtle elements of morality not denoted by conformity must be put in a vague and honorific wastebasket called rationality,

> autonomy, creativity, etc ... Frequently an effort is then made to define this wastebasket behavioristically. In such cases the autonomy category tends to become simply the opposite of the moral one, it becomes resistance to social pressure and rule just as morality was conformity to it. (1958, p. 5)

These are not only caricatures of moral stances, but they are caricatures of the wrong moral stages. We must guess that Hogan, or Kurtines and Greif, have confused the post-conventional rule-utilitarianism of stage 5 with the conventional act-utilitarianism of stage 4 (Kohlberg, 1971), and the stage 6 conception of justice and equity with the "intuitive" individualism of "stage 4 1/2" (Kohlberg, 1973b).[14] This interpretation of Hogan is supported by a recent article of his in which he brands Kohlberg as defining a "romantic" or "ideological individualism" as the end-point of development (Hogan, 1974b). It is true that stage 6 assumes the concept of rational self-interest (Morelli, 1975; Habermas, 1975; Smythe, 1974), but it is neither romantic nor individualistic. Quite to the contrary--stage 6 is built upon an ideal social selflessness that treats self and others with complete equity, and which fashions its judgments by the criterion of their potential "sharability" (Kohlberg, 1970).[15] The egoistic individualism that Hogan criticizes is transcended even by stage 5. Hogan's own moral end-point appears to describe a form of "natural law" ethics in which moral "shoulds" are derivable from pre-existing laws of natural personality or human nature.[16] This is a version of the naturalistic fallacy that Kohlberg (1971) has so

fully described. It is not surprising then that Hogan's most "mature" individuals exhibit a mixture of both rule-bound and intuitive tendencies, since this "natural law" philosophy is a mixture of just those moral positions that Hogan's scale appears to measure--stages 4 and 4 1/2. In confounding "is" and "ought," this "mature" ethic combines the rigidness of conventionality with the opportunistic idealism of relativistic thinking.

In conclusion therefore, we find that none of the three lines of counterevidence is likely to bring the Kohlberg fish to shore, and we have a sneaking suspicion that none of the three hooks is even properly baited.

KOHLBERG'S THEORY OF MORAL DEVELOPMENT

So far we have tackled the misinterpretation of research findings at their own level of analysis. There are however deeper misunderstandings of Kohlberg's theory that underlie these, a clarification of which helps us to see a certain unity to Kurtines and Greif's approach.

We find key symptoms in Kurtines and Greif's initial description of Kohlberg's theory. Along with a rather inexplicable reference to its being a projective test, they basically assimilate Kohlberg's work to the standard psychometric model. They call it the "Moral Judgment Scale," whereas Kohlberg has explicitly distinguished cognitive-developmental stage theories from psychometric scales (Kohlberg, 1968; Kohlberg & de Vries, 1971). This point has been reiterated by Elkind (1971), to some extent by Tuddenham (1971), and is well taken by Wohlwill (1973b). Despite brave and industrious attempts (Uzgiris &

Hunt, 1973; Tuddenham, 1971), cognitive-developmental theories are inappropriately converted into traditional scales for several reasons. The first of these we have already noted--their discontinuity assumptions (Beilin, 1971). Calling them ordinal or nominal scales does not work either, since theories of the Piaget or Kohlberg variety are concerned with the absolute logical qualities of within-subjects cognitive organizations, not the relative achievement of these subjects with respect to group means, expressed in terms of empirical curves of percent passing a given item. Tasks are restructured, not passed or failed. There is also much less emphasis placed upon age-relatedness in a stage-theory, since the form of development is defined independently of its rate.

These differences reflect the fact that cognitive-developmental theories deal with fundamentally different phenomena. They are concerned with the structurally invariant sequential interactions of organism and environment that govern a sense of the logic to physical and social reality. They are not concerned with the assimilation of particular cultural information. This does not mean that Kohlberg's theory cannot give birth to a measuring instrument. It does mean that instead of assimilating that instrument to the psychometric model, the latter has to accommodate to a wholly emergent theoretical system, one that defines measurement in a totally different way. This involves a reordering of priorities, in which "construct validity" becomes the dominant concern, particularly with regard to fulfilling the dual criteria of "stage" and "sequence" (Laurendeau &

Pinard, 1962).

A clear demonstration of Kurtines and Greif's lack of awareness of the paradigm difference is his criticism of the administration of the Kohlberg "scale." Because the interviewer encourages explanation of the subject's original response, adapting probe questions to that initial answer, Kurtines and Greif argue (p. 455) that the questionnaire lacks standardization, each person receiving a different set of questions. This is first of all an incomplete account of the research that has been done on Kohlberg's theory, most of which actually uses a standard questionnaire. The results achieved with this method have tallied extremely well with results from interviews, although well-probed interviews always yield more reliable and more consistently scored protocols. Secondly, Kohlberg's critics often appear not to realize that the instrument measures the form of moral reasoning (the way in which ethical issues are approached), not the specific content of the subject's response. As a result, variations in the content of administration and the effect they have on the content of answers are much less important. They do have some significance, because content and form are related in practice. But even if we ignore this for the moment, Kurtines and Greif's remarks about standardization represent the logical positivist's operationalism that has too long had a deadly grip on psychology. How can one sensibly ask the exact same questions of two different subjects? If the first subject says "Heinz should steal the drug," and the interviewer asks "Why should he?", must he then ask the same probing question of a second subject

who says "Heinz should not steal the drug"? Simply asking "Why?" does not "standardize" the probe, because we are not dealing with isolated, objective words with fixed meanings--like it or not, the subject interprets each question, and does so in the context of what has already been said in the test or interview. Such a fact of life is partly acknowledged even in the administration of "standardized" tests such as the WISC. For example, it is often necessary to "make sure the subject understands the question." Standardized questioning ignores the fact that meanings develop through interpersonal understanding in the test situation, and leads to "objective" mindlessness, and to trivial results (as pointed out by Ryan, 1972, and Hudson, 1972). It also leaves us with the possible problem that regularities in our results are entirely a function of specific factors associated with superficial characteristics of question form. We may obtain delightfully "clean" data, free from any contamination by the person who was tested and therefore totally out of contact with that person's thought. This is another example of what Riegel (1975a) has called "subject-object alienation" in the testing situation. For a more realistic and less alienated discussion of interview method, see Piaget (1929, pp. 1-32). For those who feel that the only alternative to standardized test formats is unregulated and subjective chaos, Piaget's detailed outlines of alternative methodological rules may provide some reassurance. Also, as Cowan et al. have pointed out:

> We no longer have to assume that the "unstandardized" approach works against the discovery of lawful relationships, since

> Piaget's method has led to some of the most replicable results in the field of psychology. (1969, p. 272)

The same thing applies to Kurtines and Greif's criticism of the "judgmental nature of the coding procedure" (p. 456). If coding does not involve judgment, then we regress to the same mindless mechanism just discussed. What do Kurtines and Greif's own criteria of objectivity consist in, if not particular judgments imbedded in the positivist tradition of psychometric inquiry? Cognitive psychology must face up to the fact that statements have meaning, and that it is the study of ideas, not patterns of words, that is its appropriate focus. As Levine (1974) has pointed out, we are all too ready to shirk responsibility and forget that human intelligence is the basic research instrument, not experimental design or psychometric technique. From Kurtines and Greif's point of view, human judgment is clearly to be equated with error, with bias. This is surely not the case when we have judgments that are ordered by generalizable, rational criteria, within explicitly defined categories, dictated by a clearly stated and self-consistent theory. However much effort it may require to obtain the scoring manual from Kohlberg, one finds that, although there is clearly room for improvement, it is fairly exemplary in the detailed constraints it places upon coding judgments, ensuring that they are as judgmental as possible in precisely the most theoretically justifiable way. To eliminate judgment from coding procedures brings with it the implication that we must eliminate judgment from research in general. It is not the presence,

but the absence of a judgmental orientation that defines subjectivity (Cassirer, 1923). Kohlberg should therefore be exhorted to iron out the inadequacies of the scoring method, or develop a more sophisticated method, rather than to throw out the scoring method on the groupds that it requires human intelligence and understanding for its correct operation.

It is precisely the development of such detailed judgmental criteria, elaborated and constantly revised in the light of longitudinal studies of natural moral development, and securely located in the context of a well-grounded and thoroughly argued philosophic ethical theory, that are perhaps responsible for the irritating delays in publication that Kurtines and Greif rightly point out. It is only when such construction of a theoretical-empirical edifice is reduced to a matter of reliability and validity in a psychometric scale that Kohlberg's programme of work as a whole becomes "inexplicable." Kurtines and Greif's criticisms, often bringing out problems that need to be dealt with, would only become forceful if they were directed at the same time towards Kohlberg's stages as a theory of moral development. One cannot attack such a theory without bravely standing astride the unfortunate gulf that has been forged between empirical technology and its theoretical focus.[17] In our present instance, this would require paying some attention to at least two kinds of characteristics of Kohlberg's cognitive-developmental thesis. The first is its theoretical presuppositions: equilibration theory, structural psychology of cognition, subject-object interactionism, deontological view of

morality as justice rather than welfare, etc. We must ask, are these justifiable, self-consistent, coherent and correctly interpreted within the theory? This has by no means gone undisputed (Alston, 1971; Peters, 1971; Siegal, 1975). The second is the internal logic of the theory, the conceptual organization of interview material through stage-by-stage differentiation and integration of categories and relations. The most fundamental question about a stage theory must be not the Aristotelian one of whether stages just reflect empirical regularities, but the Galileian one of whether or not their descriptions make sense as a lawful sequential ordering of progressively advancing and more adequate cognitive organizations. Several writers have found lacking the philosophical grounds of Kohlberg's treatment of post-conventional morality (Habermas, 1975; Buck-Morss, in press; Morelli, 1975; Gibbs, 1975; Siegel, 1974; Siegal, 1975; Baumrind, 1975; Smythe, 1974).[18] The paucity of data on the post-conventional stages has meant that such discussion has been confined to the level of philosophical presuppositions. Unfortunately, the question of the adequacy of the psychological theory cannot be answered without adding to this empirical recourse to detailed interview material, since it is concrete moral judgments that such a hermeneutic system claims to interpret and evaluate as more or less adequate.[19]

Kurtines and Greif on the other hand exhibit the opposite shortcoming. Although they show some concern for the explanatory adequacy of the theory, when they do not take the theoretical logic and

presuppositions of Kohlberg's framework into account, but interpret it merely as a psychometric scale, they confuse the following: values with morality, conformity with moral development, moral style with moral judgment, moral statements with moral structure, the content of moral choice with the form of the reasoning behind it, specific behaviors with moral action, psychometric scale with stage theory, scale types with stages, reliability/homogeneity with internal consistency of stages, standard error of measurement with horizontal decalage, measurement stability with structural stability, psychometric validity with stage sequence, and so on.

The result is, inadvertently, an expression of disappointment that Kohlberg is not in the projective test business. From this point of view it must seem sadly inexplicable that Kohlberg is weighed down with so much theory and so much laborious longitudinal research. A more helpful analysis would detail and criticize the ways in which the cognitive-developmental theory of morality succeeds or fails in the attempt to provide a radically different perspective on socialization, and a new paradigm for structural research, grounded in a structuralist rather than a logical positivist epistemology. The empirical utility of the cognitive-developmental approach has surely been demonstrated, but the question of its truth has yet to be seriously confronted. It is not, as Kurtines and Greif suggest, Kohlberg who discourages independent research, but the very conceptual and empirical difficulty of the task itself. To reject Kohlberg's perspective prematurely on the basis of criteria from another domain can only be a way of avoiding this important task.

References

Alston, W. P. Comments on Kohlberg's "From is to Ought." In T. Mischel (Ed.). Cognitive Development and Genetic Epistemology. New York: Academic Press, 1971.

Arbuthnot, J. Modification of moral judgment through role-playing. Developmental Psychology, 1975, 11, 3, 319-324.

Bakan, D. The test of significance in psychological research. Psychol. Bull., 1966, 66, 6, 423.

Bandura, A. Social learning of moral judgments. J. Pers. Soc. Psychol., 1969, 11, 3, 275-279.

Baumrind, D. It neither is nor ought to be. In E. Kennedy (Ed.). Psychology and Ethics. Chicago: Loyola University, 1975.

Beilin, H. Developmental stages and developmental processes. In D. R. Green, M. P. Ford, & G. B. Flamer (Eds.). Measurement and Piaget. New York: McGraw-Hill, 1971.

Blasi, A. Role-taking and the development of social cognition. Paper presented at the annual meeting of the American Psychological Association, Chicago, August 1975.

Blatt, M. Studies on the effects of classroom discussion upon children's moral development. J. Moral Educ., 1975, 42, 129-161.

Boyd, D. From conventional to principled morality. Unpublished manuscript, Harvard University, 1972.

Broughton, J. M. "College regression" in moral development. Unpublished manuscript, Harvard University, 1970.

Broughton, J. M. The development of natural epistemology in adolescence. Unpublished doctoral dissertation, Harvard University,

1974.

Broughton, J. M. The cognitive-developmental approach to "epistemology" and its relationship with logical and moral stages. Paper presented at the bi-annual meeting of the Society for Research in Child Development, Denver, April 1975.

Brown, R., & Herrnstein, R. Psychology. Boston: Little, 1974.

Buck-Morss, S. Socio-economic bias in Piaget's theory and its implications for cross-cultural studies. Human Development, in press.

Byrne, D. Role-taking development in the years from ten to adulthood. Unpublished doctoral dissertation, University of Rhode Island, 1973.

Cassirer, E. Substance and Function. New York: Dover, 1923.

Cattell, R. B. Validity and reliability: a proposed more basic set of concepts. J. Educ. Psychol., 1964, 55, 1-22.

Chomsky, N. Review of B. F. Skinner's "Verbal Behavior." Language, 1959, 35, 26-58.

Chomsky, N. Language and Mind. New York: Harcourt, Brace and World, 1968.

Cohen, J. The statistical power of abnormal/social psychological research. J. Abnorm. Soc. Psychol., 1962, 65, 145.

Cohen, J. Some statistical issues in psychological research. In B. Wolman (Ed.). Handbook of Clinical Psychology. New York: McGraw-Hill, 1966.

Costa, P. Development of the ego-ideal. Unpublished manuscript, Harvard University, 1971.

Cowan, P. A., Langer, J., Heavenrich, J., & Nathanson, M. Social learning and Piaget's cognitive theory of moral development. J. Pers. Soc. Psychol., 1969, 11, 261-274.

Cronbach, L. J., & Gleser, G. C. Assessing similarity between profiles. Psychol. Bull., 1953, 50, 456-473.

Dewey, J. Human Nature and Conduct. New York: Henry Holt, 1922.

Dworkin, E. S. The effects of imitation, reinforcement, and cognitive information on the moral judgments of children. Unpublished doctoral dissertation, University of Rochester, 1968.

Elkind, D. Conservation and concept formation. In D. Elkind and J. H. Flavell (Eds.). Studies in Cognitive Development. New York: Oxford University Press, 1969.

Elkind, D. Two approaches to intelligence: Piagetian and psychometric. In D. R. Green, M. P. Ford, and G. B. Flamer (Eds.). Measurement and Piaget. New York: McGraw-Hill, 1971.

Flavell, J. H., & Wohlwill, J. Formal and functional aspects of cognitive development. In D. Elkind and J. H. Flavell (Eds.). Studies in Cognitive Development. New York: Oxford University Press, 1969.

Fowler, J. Toward a developmental perspective on faith. Relig. Educ., 1974, 69, 2, 207-219.

Gibbs, J. C. The Piagetian approach to moral development: an overview. In J. Meyer, B. Burnham, & J. Chilvar (Eds.). Values Education: Theory/Practice/Problems/Prospects. Waterloo, Ontario: Wilfrid Laurier University Press, 1975.

Gibbs, J. C., & Fedoruk, C. J. Experimentally induced change in moral

judgment. Paper presented at the symposium of the Jean Piaget Society, Philadelphia, June 1975.

Gilligan, C. The development of moral reasoning about sexual dilemmas. Paper presented to the Sex Information Council of the United States Conference, December 1971.

Gilligan, C., & Kohlberg, L. From adolescence to adulthood: the rediscovery of reality in a post-conventional world. Paper given at the annual meeting of the Jean Piaget Society, June, 1974.

Haan, N., Smith, M. B., & Block, J. Moral reasoning of young adults: political social behavior, family background, and personality correlates. J. Pers. Soc. Psychol., 1968, 10, 185-201.

Hartshorne, H., & May, M. Studies in Character. New York: Macmillan, 1929.

Hamlyn, D. W. Conditioning and behavior. In R. Borger and F. Cioffi (Eds.). Explanation in the Behavioral Sciences. Cambridge, England: Cambridge University Press, 1970.

Hickey, J. The effects of guided moral discussion upon youthful offenders' level of moral judgment. Unpublished doctoral dissertation, Boston University, 1972.

Hogan, R. Moral conduct and moral character: a psychological perspective. Psychol. Bull., 1973, 79, 4, 217-232.

Hogan, R. On theoretical egocentrism and the problem of compliance. American Psychologist 1974, 30, 5, 533-540.

Holstein, C. B. The relation of children's moral judgment level to that of their parents and to communication patterns in the family. In R. C. Smart and M. S. Smart (Eds.). Readings in Child

Development and Relationships. New York: Macmillan, 1972.

Holstein, C. B. A longitudinal study of family influences on adolescent moral judgment development. Paper presented at the biennial meeting of S.R.C.D., Philadelphia, March-April, 1973.

Horkheimer, M. The Eclipse of Reason. New York: Seabury Press, 1974.

Hudson, L. The Cult of the Fact. New York: Harper, 1972.

Kaplan, B. Meditations on genesis. Human Development, 1967, 10, 65-87.

Keasey, C. B. Experimentally induced changes in moral opinions and reasoning. Journal of Personality and Social Psychology, 1973, 26, 1, 30-38.

Kohlberg, L. The development of modes of moral thinking and choice in the years ten to sixteen. Unpublished doctoral dissertation, University of Chicago, 1958.

Kohlberg, L. The development of children's orientations toward a moral order: sequence in the development of moral thought. Vita Humana, 1963, 6, 11-33 (a).

Kohlberg, L. Moral development and identification. In H. Stevenson (Ed.). Child Psychology. Yearbook of Nat. Soc. for the Study of Education. Chicago: University of Chicago Press, 1963(b).

Kohlberg, L. Early education: a cognitive-developmental approach. Child Development, 1968, 39, 1013-1062.

Kohlberg, L. Stage and sequence: the cognitive-developmental approach to socialization. In D. A. Goslin (Ed.). Handbook of

Socialization Theory and Research. Chicago: Rand McNally, 1969(a).

Kohlberg, L. The relations between moral judgment and moral action. Paper presented at the Institute of Human Development, University of California, Berkeley, March 1969(b).

Kohlberg, L. Mimeo scoring manual. Harvard Graduate School of Education, 1969(c).

Kohlberg, L. Education for justice: a modern statement of the Platonic view. In T. Sizer (Ed.). Moral Education. Cambridge, Mass.: Harvard University Press, 1970.

Kohlberg, L. From is to ought: how to commit the naturalistic fallacy and get away with it in the study of moral development. In T. Mischel (Ed.). Cognitive Development and Genetic Epistemology. New York: Academic Press, 1971.

Kohlberg, L. Continuities and discontinuities in childhood and adult moral development revisited. In P. B. Baltes and K. W. Schaie (Eds.). Lifespan Developmental Psychology--Personality and Socialization. New York: Academic Press, 1973(a).

Kohlberg, L. The claim to moral adequacy of a highest stage of moral judgment. Journal of Philosophy, 1973, 70, 18, 630-646 (b).

Kohlberg, L. Moralization: the cognitive-developmental approach. In T. Lickona (Ed.). Morality: Theory, Research and Social Issues. Holt, Rinehart and Winston, 1975(a).

Kohlberg, L. Continuities and discontinuities revisited again.

Talk given at a colloquium on "Dialectics and Development," Wayne State University, Detroit, February 1975(b).

Kohlberg, L. Moral stage scoring manual. Mimeo, Center for Moral Education, Harvard, June 1975(c).

Kohlberg, L. Response to critics. Paper presented at the meeting of the Society for Research in Child Development, Denver, April 1975(d).

Kohlberg, L., & de Vries, R. Relations between Piaget and psychometric assessments of intelligence. In C. Lavatelli (Ed.). The Natural Curriculum. Urbana, Illinois: E.R.I.C., 1971.

Kohlberg, L., & Elfinbein, D. The development of moral judgments concerning capital punishment. American Journal of Orthopsychiatry, in press.

Kohlberg, L., & Freundlich, D. Moral judgment in youthful offenders. In L. Kohlberg and E. Turiel (Eds.). Moralization, the Cognitive-Developmental Approach. New York: Holt, Rinehart and Winston, in press.

Kohlberg, L., & Kramer, R. Continuities and discontinuities in childhood and adult moral development. Human Development, 1969, 12, 93-120.

Krebs, R. L. Some relationships between moral judgment, attention and resistance to temptation. Unpublished doctoral dissertation, University of Chicago, 1967.

Krebs, D., & Rosenwald, A. Moral reasoning and moral behavior in conventional adults. Unpublished manuscript, Harvard University,

1973.

Kuhn, D., & Langer, J. Cognitive-developmental determinants of imitation. Unpublished manuscript, University of California, Berkeley, 1968.

Kurtines, W., & Greif, E. B. The development of moral thought: review and evaluation of Kohlberg's approach. Psychol. Bull., 1974, 81, 8, 453-470.

Laurendeau, M., & Pinard, A. Causal Thinking in the Child. New York: International Universities Press, 1962.

LeFurgy, W., & Woloshin, G. W. Immediate and long-term effects of experimentally induced social influence in the modification of adolescents' moral judgments. J. Pers. Soc. Psychol., 1969, 12, 104-110.

Levine, M. Scientific method and the adversary model. American Psychologist, 1974, 29, 661-677.

Lewin, K. The conflict between Aristotelian and Galileian modes of thought in contemporary psychology. J. Gen. Psychol., 1931, 5, 141-177.

Lichtman, R. Symbolic interactionism and social reality: some Marxist queries. Berkeley Journal of Sociology, 1970, 10, 75-94.

Lieberman, M. Calculation of a polychoric correlation coefficient. Paper presented at the Spring meeting of the Psychometric Society, Princeton, New Jersey, 1969.

Lieberman, M. Psychometric analysis of developmental stage data.

Paper presented at the Annual Meeting of the American Psychological Association, Montreal, August 1973.

Loevinger, J. Recent research on ego development. Paper presented at bi-annual meeting of the Society for Research in Child Development, Philadelphia, March 1973.

Loevinger, J. Issues in the measurement of moral development. In W. H. Holtzman (Ed.). Moral Development. Publications of the International Conference on Testing Problems, Educational Testing Service, Princeton, 1974.

Loevinger, J. Ego Development: Conceptions and Theories, in press.

Loevinger, J., & Wessler, R. Measuring Ego Development, Vol. 1. San Francisco: Jossey-Bass, 1970.

Merleau-Ponty, M. The Structure of Behavior. Boston: Beacon Press, 1963.

Milgram, S. Obedience to Authority. New York: Harper and Row, 1974.

Morelli, E. A dialectical critique of Kohlberg's moral development theory. Unpublished manuscript, Department of Philosophy, University of Toronto, 1975.

Perry, W. Forms of Intellectual and Ethical Development in the College Years. Cambridge, Mass.: Harvard Bureau of Study Counsel, 1968.

Peters, R. S. Moral development: a plea for pluralism. In T. Mischel (Ed.). Cognitive Development and Genetic Epistemology. New York: Academic Press, 1971.

Piaget, J. The Child's Conception of the World. New York: Harcourt, Brace, 1929.

Piaget, J. The general problems of the psychobiological development of the child. In J. M. Tanner and B. Inhelder (Eds.). Discussions on Child Development, Vol. 4. New York: International Universities Press, 1969.

Pinard, A., & Laurendeau, M. "Stage" in Piaget's cognitive-developmental theory: exegesis of a concept. In D. Elkind and J. H. Flavell (Eds.). Studies in Cognitive Development. New York: Oxford University Press, 1969.

Porter, N., & Taylor, N. How to assess the moral reasoning of students: A teacher's guide to the use of Lawrence Kohlberg's stage-developmental method. Profiles of Practical Education, No. 8. Toronto: The Ontario Institute for Studies in Education, 1972.

Rawls, J. A Theory of Justice. Cambridge, Harvard University Press, 1971.

Rest, J. R. Developmental psychology as a guide to value education: a review of "Kohlbergian" programs. Review of Educational Research, March 1974, in press.

Rest, J. R. Bibliography on development of moral-political judgment. Mimeo, Institute of Child Development, University of Minnesota, 1975.

Rest, J. R., Turiel, E., & Kohlberg, L. Level of moral development as a determinant of preference and comprehension of moral judgments

made by others. J. Personality, 1969, 37, 225-252.

Rex, J. Max Weber, value relevance, scientific law and ideal types. Paper presented at the seventh annual meeting of Cheiron. The International Society for the History of the Behavioral and Social Sciences, Ottawa, June 1975.

Riegel, K. F. From traits and equilibrium toward developmental dialectics. Paper presented at the Nebraska Symposium on Motivation, 1974.

Riegel, K. F. Subject-object alienation in psychological experimentation. In K. F. Riegel (Ed.). The Development of Dialectical Operations. Basel: Karger, 1975(a). Also Human Development, 1975, 18, in press.

Riegel, K. F. Dialectical psychology and cognitive development. Talk given at a colloquium on "Dialectics and Development," Wayne State University, Detroit, February 1975(b).

Ryan, J. IQ--the illusion of objectivity. In K. Richardson and D. Spears (Eds.). Race and Intelligence. London: Penguin.

Saltzstein, H. D., Diamond, R. M., & Belenky, M. Moral judgment level and conformity behavior. Developmental Psychology, 1972, 7, 327-336.

Sanders, S. L., & Cone, D. M. A critique of Kohlberg's theory of moral development from the viewpoint of interbehavioral psychology. Interbehavioral Quarterly, 1975, 6, 2, 8-22.

Selman, R., & Byrne, D. Manual for scoring social role-taking stages in moral and social dilemmas. Unpublished manuscript,

Harvard University, 1973.

Siegal, M. Learning moral competence. Unpublished paper, Oxford University, 1975.

Siegel, H. Kohlberg: evaluative status and moral adequacy. Unpublished manuscript, Department of Philosophy, Harvard University, 1974.

Simpson, E. L. Moral development research: a case study of scientific cultural bias. Human Development, 1974, 17, 81-106.

Smedslund, J. The acquisition of the conservation of substance and weight in children: III extinction of conservation of weight acquired "normally" and by means of empirical controls on a balance. Scandinavian Journal of Psychology, 1961, 2, 85-87.

Smith, J., & Firestone, I. The development of moral judgment from childhood through maturity: a new social scale. Unpublished manuscript, Wayne State University, 1973.

Smythe, O. Kohlbergian moral development theory and the legacy of John Dewey. Unpublished manuscript, Department of Philosophy, Harvard University, 1974.

Sullivan, E. V., & Quarter, J. Psychological correlates of certain post-conventional moral types: a perspective on hybrid types. Journal of Personality, 1972, 40, 2, 149-161.

Sutton, L. E. An investigation of invariant sequence in a cognitive-developmental theory of moral development. Unpublished master's thesis, Connecticut College, 1972.

Taylor, C. The Explanation of Behavior. London: Routledge and

Kegan Paul, 1964.

Tracy, J. J., & Cross, H. J. Antecedents of shift in moral judgment. Journal of Personality and Social Psychology, 1973, 26, 2, 238-244.

Troutner, L. The confrontation between experimentalism and existentialism: from Dewey through Heidegger and beyond. Harvard Educational Review, 1969, 39, 1, 124-154.

Tuddenham, R. Theoretical regularities and individual idiosyncrasies. In D. R. Green, M. P. Ford, and G. B. Flamer (Eds.). Piaget and Measurement. New York: McGraw-Hill, 1971.

Turiel, E. An experimental test of the sequentiality of developmental stages in the child's moral judgments. J. Pers. Soc. Psychol., 1966, 3, 611-618.

Turiel, E. Developmental processes in the child's moral thinking. In P. Mussen, J. Langer, and M. Covington (Eds.). Trends and Issues in Development Psychology. New York: Holt, Rinehart and Winston, 1969.

Turiel, E. Stage transition in moral development. In R. M. Travers (Ed.). Second Handbook of Research on Teaching. Chicago: Rand McNally, 1972.

Turiel, E. Conflict and transition in adolescent moral development. Child Development, 1974, 45, 14-29 (a).

Turiel, E. A comparative analysis of moral knowledge and moral judgment in males and females. Unpublished manuscript, University of California, Santa Cruz, 1974(b).

Turiel, E., Kohlberg, L., & Edwards, C. Cross-cultural studies in moral development. In Kohlberg, L. and Turiel, E. (Eds.). Moralization: the cognitive-developmental approach. New York: Holt, Rinehart and Winston, in press.

Turner, E. In Kuhlen, M. G. and Thompson, G. G. (Eds.). Psychological Studies in Human Development. New York: Appleton-Century Crofts, 1952.

Ullian, D. A cognitive-developmental theory of sex-role attitudes. Unpublished doctoral dissertation, Harvard University, 1974.

Uzgiris, I., & Hunt, J. McV. Toward ordinal scales of psychological development in infancy. Unpublished manuscript, Clark University, 1972.

van den Daele, L. A developmental study of the ego-ideal. Genetic Psychology Monographs, 1968, 78, 191-256.

Weber, M. The Methodology of the Social Sciences. In E. A. Shils and H. A. Finch (Eds.). Glascoe, Illinois: The Free Press, 1949.

Weisbrodt, S. Moral judgment, sex and parental identification in adults. Developmental Psychology, 1972, 2, 396-402.

Westermarck, E. Ethical Relativity. London: Routledge and Kegan Paul, 1932.

Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill, 1962.

Wohlwill, J. Methodology and research strategy in the study of developmental change. In L. R. Goulet and P. B. Baltes (Eds.). Lifespan Developmental Psychology. New York: Academic Press,

1970.

Wohlwill, J. The Study of Behavior Development. New York: Academic Press, 1973.

Zinn, V. Methodological issues in the study of the development of moral judgment. Unpublished manuscript, Clark University, 1970.

Footnotes

[1]Thanks to the following for their critical suggestions: Augusto Blasi, Daniel Candee, John Gibbs, Alan Gordon, Lawrence Kohlberg, Marc Lieberman, Jane Loevinger, Robert Selman, and Elliot Turiel.

[2]Although it appears less than stringent, the 50% criterion claims significance because this is, for all stages, about two standard deviations above the total population mean percentage (Kohlberg, 1958).

[3]It could be added here that the second factor (indicating situational specificity of response) was not significant.

[4]Kurtines and Greif engage in several apparent misrepresentations with regard to these empirical indications of consistency. With respect to the first one, they twice suggest as examples distributions of an individual's score across stages that do not fit Kohlberg's recent reports. In one instance (p. 455) they give us a sample profile of an individual's reasoning as "40% stage 1, 30% stage 2, 20% stage 3, 10% stage 4." In a second instance, (p. 457) they say "For example, a subject whose modal stage is 5 may have 45% of his responses at this stage, with 35% of his responses at stage 4, and 20% at stage 6." They further suggest rather mischievously (p. 457) that given the stage score of a subject as stage 3, "one does not know whether this is $3 \pm .2$, 3 ± 1.2, or even 3 ± 3.0." The second example is particularly misleading, since stage 5 tends to be one of the "purer" stages, with the least admixture

of adjacent stages (Kohlberg, 1973a). All three instances misrepresent Kohlberg's findings by creating an impression of more spread in profile than is generally found.

[5]This cognitive consistency across stories is confirmed in Lieberman's (1973) further analysis of moral judgment data. Lieberman points to the possibility that different stories equally tap the typology, but in different ways. This more interesting approach to lack of cross-situational generality asks how it reflects the differential difficulty of stories, and the differential power that they have for discriminating the various stages. Lieberman frames his approach in a way that is theoretically compatible with the Piagetian notion of "horizontal decalage" that describes the ordering within cross-task variability in performance (Flavell & Wohlwill, 1969; Pinard & Laurendeau, 1969). According to his account, different patterns of performance across the different stories are to be expected at different points in the developmental sequence, since each story presents a peculiar constellation of conflicted moral issues. Some conflicts discriminate some stages better than others, because of the unique internal logic of each stage. Thus preserving the story-by-story profile of moral judgment levels yields information about position in the developmental sequence that is lost in simply summing to a profile of percentages as Kohlberg has done.

One thing that Lieberman's work draws our attention to is that there is a certain essential lack of item homogeneity in the moral judgment questionnaire. Where Kohlberg assumes item homogeneity in

simply summing percentages of reasoning at each stage, Lieberman capitalizes on whatever item heterogeneity there is. In so doing he bolsters Cattell's warning against the kind of demand for high homogeneity that Kurtines and Greif make, since:

> A good test, like a good engine, may perform its function best when its parts are differentiated. Whether water is better raised by a homogeneous sponge or a nonhomogeneous pump depends on the specific purpose. The aim of a good test design is surely a properly functioning single instrument, and in many cases the optimal internal consistency for such an instrument is one which is made to fall below a certain homogeneity coefficient value. (Cattell, 1964, p. 1)

While internal consistency must be demonstrated in subjects' thought for Kohlberg to substantiate his claims about stages, internal consistency in the sense of item homogeneity in the questionnaire is not a hard and fast requirement. As we have seen, different situations may be harder at some stages than others, and this is not only to be expected as a reflection of "decalage" in development (Pinard & Laurendeau, 1969) but it is also a sign that the questionnaire as a whole is a "properly functioning single instrument" reflecting the developmental course of moral reasoning.

[6]A standardised version of the scoring system has recently become available (Kohlberg, 1975c).

[7]A separate, but important, issue is that any remaining lack of consistency in the stage attributed to subjects' reasoning may either

be a function of lack of "stageness" in this cognitive domain, or a function of inadvertent overlap in the descriptions of the stages in the scoring manual (Broughton, 1974). Thus within-subjects variability in stage score is ambiguous. Its ambiguity spreads to a theoretical level as well, since where there is relatively high "stage mixture," it is not clear whether this is to be taken as a sign of snags in the scoring system, or as objective evidence of stage transition. Dissatisfaction has been voiced with the failure to characterize transitional states per se, rather than in terms of stage mixture (Sullivan & Quarter, 1972; Turiel, 1974a). Kohlberg himself (1975b) has recently become more concerned with this issue, although he has not yet reached the stage of providing an antidote.

[8]Loevinger (Chapter 9, pp. 29-30) correctly points out that Kohlberg's use of a stage profile is often confused and leads to scoring ambiguity.

[9]A possible objection is that this reinterpretation is merely pragmatic "saving the phenomena"--a change of tune to preserve intact the assumption of regular stage-by-stage progress in development. However, this distinction of naive and relativistic egoism is substantiated for morality by Sullivan and Quarter (1972), and in several other areas of cognitive-developmental research (van den Daele, 1968; Loevinger, 1970; Gilligan, 1971; Byrne, 1973; Selman & Byrne, 1973; Fowler, 1974; Ullian, 1974; Broughton, 1974). Furthermore, the justification for this interpretation is logically and theoretically more firmly grounded than the Kohlberg and Kramer

rationale, both within the narrower moral domain and within the broader realm of cognitive-developmental theory (Turiel, 1974a).

[10]A recent report by Gibbs and Fedoruk (1975) has suggested that the nature, and the demonstration, of transition from conventional to post-conventional morality may be qualitatively more difficult than for other transitions. A continuation of Boyd's (1972) work supports this misgiving (Boyd, personal communication).

[11]For help in understanding the critique of his work I am indepted to personal communications from Turiel.

[12]Kohlberg et al. did not help matters in the early days by their custom of converting stage score profiles into a continuous dimension called the "moral maturity score."

[13]A recent study by Turiel (1974b) indicates that the 1972 Holstein study may be of limited generalizability. In this study, Turiel found that sex differences in moral stage attainments of 10-16 year-old girls and boys were not significant. Also, Turiel reports a study which claims that sex differences disappear when level of education is taken into account (Weisbrodt, 1970).

[14]The extent of Hogan's misunderstanding is witnessed by his equation of moral principles with learned rules for self-control (Hogan, 1973, p. 220).

[15]It is the "social" interpretation of morality, not the individualistic one, that embarrasses Kohlberg's theory, since the latter has difficulty in accounting for people's obligations to <u>themselves</u> (Morelli, 1975; Smythe, 1974; see also Blasi, 1975).

[16]Hogan (1973, p. 219) explicitly denies the distinction between morality and social rules, equating the moral with etiquette and courtesy. For him, "moral conduct is fundamentally irrational, that differences in even such obviously cognitive phenomena as moral judgments derive from more basic personological structures" (1973, p. 225).

[17]Kohlberg (1975d) has recently discussed a reconciliation of measurement theory with structural-developmentalism, and pointed out how the former is to be reinterpreted in the light of the latter (particularly with regard to the notion of "construct validity").

[18]Cf. Troutner, 1969; Lichtman, 1970; and Blasi, 1975, regarding the possible intellectual-historical origins of these philosophical problems.

[19]As Riegel (1975b) has indicated, Kohlberg has tended to succumb to a use of the "stage" notion that has encouraged a taxonomic attitude towards development. A main source of misunderstanding is that it is no longer clear exactly what kinds of concrete reasoning the stages refer to, what moral concepts are, or what exactly "structures" of reasoning imply about the relationships between these concepts. As a result, Kohlberg has unwittingly obscured what empirical claims are actually being made.